

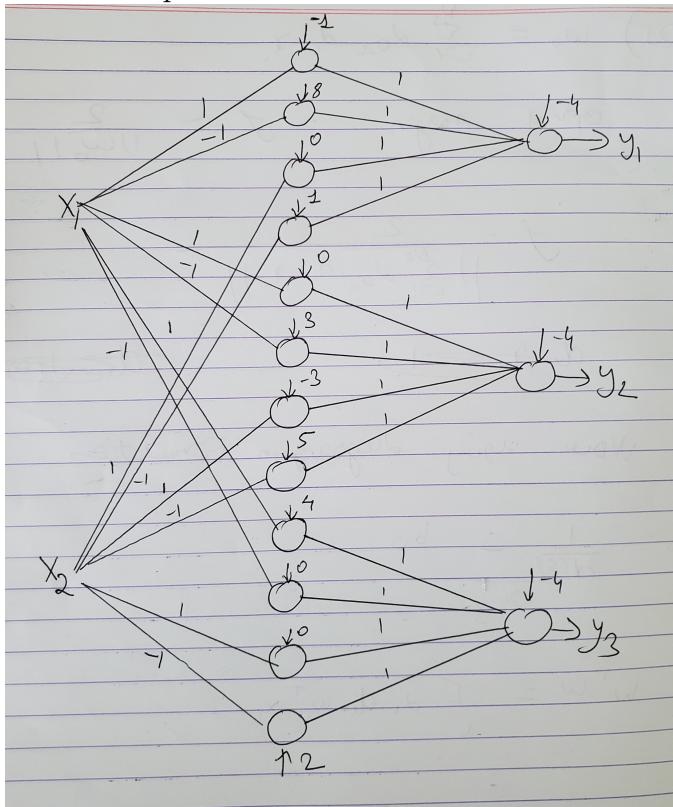
Homework Assignment 4

Rushabh Ashok Dharia
CSCI-B555 Machine Learning

April 27, 2019

Note: Homework was discussed with Varun Miranda

Ans 1. Explanation



As mentioned in the question we need to design a Neural Network with 3 output units we can consider that Class 4 is recognized when $y_i = -1$ for $1 \leq i \leq 3$

This means that when all neurons in the last layer do not fire (or output -1) we can say that the point belongs to class 4. This can be easily showed by adding another layer which will take the outputs y_1 , y_2 and y_3 as inputs

Consider the first 4 neurons in the 1st Layer

1st Neuron will output 1 if $x_1 \geq 1$, otherwise 0

2nd Neuron will output 1 if $x_1 \leq 8$, otherwise 0

3rd Neuron will output 1 if $x_2 \geq 0$, otherwise 0

4th Neuron will output 1 if $x_2 \leq 1$, otherwise 0

Now, the 1st Neuron in the 2nd Layer will output 1 if all 4 neurons mentioned above output 1, else it will output 0

Similarly, the other 2 classes can be determined

Ans 2. a).

$$\begin{aligned} J(u_j) &= \sum_{j=1}^K \sum_{i=1}^N w_{ij} \|x_i - u_j\|^2 \\ &= \sum_{j=1}^K \sum_{i=1}^N w_{ij} (x_i^2 + u_j^2 - 2x_i u_j) \\ &= \sum_{j=1}^K \sum_{i=1}^N w_{ij} x_i^2 + w_{ij} u_j^2 - 2w_{ij} x_i u_j \end{aligned}$$

Taking derivative w.r.t u_j and setting it equal to 0

$$\begin{aligned} \frac{dJ(u_j)}{du_j} &= \sum_{j=1}^K \sum_{i=1}^N 0 + 2w_{ij} u_j - 2w_{ij} x_i = 0 \\ \sum_{j=1}^K \sum_{i=1}^N w_{ij} u_j &= \sum_{j=1}^K \sum_{i=1}^N w_{ij} x_i \\ \text{Therefore, } \hat{u}_j &= \frac{\sum_{i=1}^N w_{ij} x_i}{\sum_{i=1}^N w_{ij}} \end{aligned}$$

□

b). Numerator denotes the points in each cluster and the denominator denotes the number of points in each cluster. The fraction as a whole denotes the weighted average of all points in the cluster

□

Ans 3. a).

$$\begin{aligned}
y(i) &= \sum_{j=1}^K w_j(n) e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \\
\frac{\partial y(i)}{\partial w_j(n)} &= \sum_{j=1}^K e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \\
\frac{\partial y(i)}{\partial u_j(n)} &= w_j(n) e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \cdot \frac{x_i - u_j(n)}{\sigma^2(n)} \\
\frac{\partial y(i)}{\partial \sigma(n)} &= w_j(n) e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \cdot \frac{(x_i - u_j(n))^2}{\sigma^3(n)} \\
\text{Now, } E &= \frac{1}{2} \sum_{i=1}^N (d(i) - y(i))^2 \\
\frac{\partial E}{\partial w_j(n)} &= - \sum_{i=1}^N (d(i) - y(i)) \cdot \frac{\partial y(i)}{\partial w_j(n)} \\
&= - \sum_{i=1}^N (d(i) - y(i)) \sum_{j=1}^K e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \\
\frac{\partial E}{\partial u_j(n)} &= - \sum_{i=1}^N (d(i) - y(i)) \cdot \frac{\partial y(i)}{\partial u_j(n)} \\
&= - \sum_{i=1}^N (d(i) - y(i)) \sum_{j=1}^K w_j(n) e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \cdot \frac{x_i - u_j(n)}{\sigma^2(n)} \\
\frac{\partial E}{\partial \sigma(n)} &= - \sum_{i=1}^N (d(i) - y(i)) \cdot \frac{\partial y(i)}{\partial \sigma(n)} \\
&= - \sum_{i=1}^N (d(i) - y(i)) \sum_{j=1}^K w_j(n) e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \cdot \frac{\|x_i - u_j(n)\|^2}{\sigma^3(n)}
\end{aligned}$$

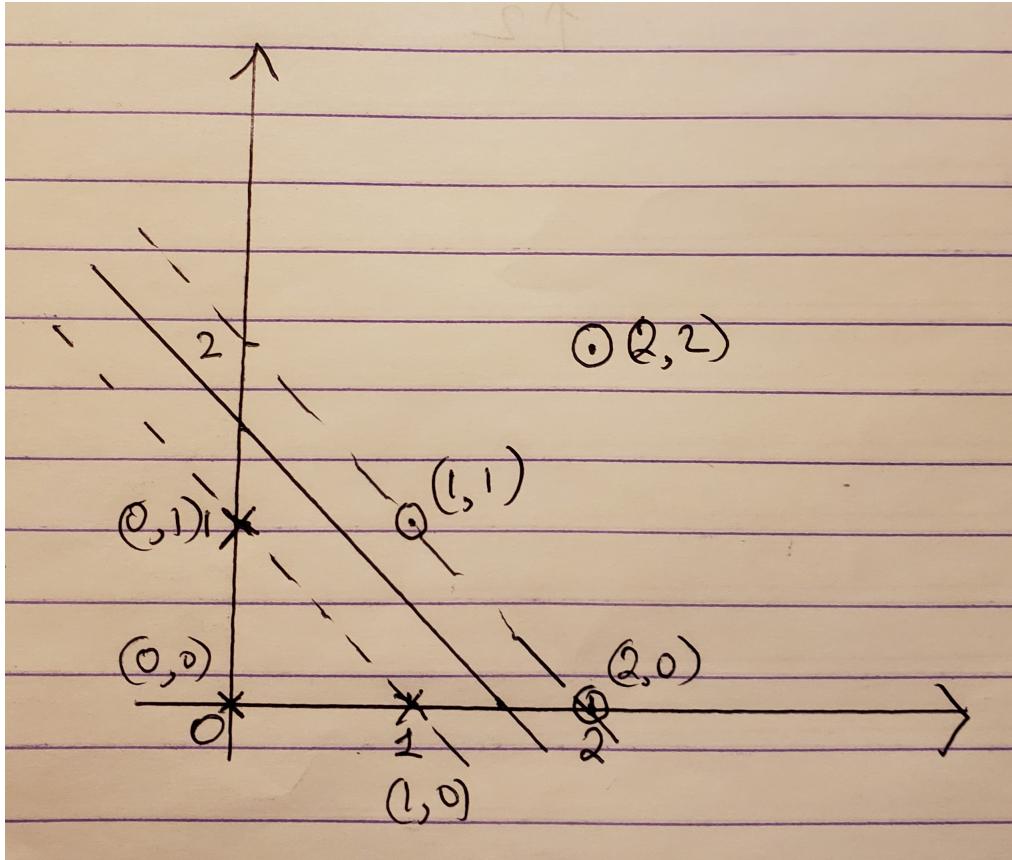
□

b).

$$\begin{aligned}
 w_{kj}(n+1) &= w_{kj}(n) - \eta_w \frac{\partial E}{\partial w_{kj}} \\
 &= w_{kj}(n) + \eta_w \sum_{i=1}^N (d(i) - y(i)) \sum_{j=1}^K e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \\
 w_{kj}(n+1) &= w_{kj}(n) - \eta_u \frac{\partial E}{\partial u} \\
 &= w_{kj}(n) + \eta_u \sum_{i=1}^N (d(i) - y(i)) \sum_{j=1}^K w_j(n) e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \cdot \frac{x_i - u_j(n)}{\sigma^2(n)} \\
 w_{kj}(n+1) &= w_{kj}(n) - \eta_\sigma \frac{\partial E}{\partial \sigma} \\
 &= w_{kj}(n) - \eta_\sigma \sum_{i=1}^N (d(i) - y(i)) \sum_{j=1}^K w_j(n) e^{\frac{\|x_i - u_j(n)\|^2}{2\sigma^2(n)}} \cdot \frac{\|x_i - u_j(n)\|^2}{\sigma^3(n)}
 \end{aligned}$$

□

Ans 4. a). Plot



□

$b)$.

The Support Vectors are $(0,1)$, $(1,0)$, $(1,1)$ and $(2,0)$

□

Ans 5.

$$J(w, b, \alpha) = \frac{w^T w}{2} - \sum_i^N \alpha_i [d_i(w^T x + b) - 1]$$

$$\phi(w) = \frac{1}{2} w^T w$$

Now, $\phi(w) = J(w, b, \alpha)$

$$\text{Therefore, } \sum_i^N \alpha_i [d_i(w^T x + b_0) - 1] = 0$$

$$\sum_i^N \alpha_i d_i w^T x + \sum_i^N \alpha_i d_i b_0 - \sum_i^N \alpha_i = 0 \quad (1)$$

$$\text{Now, } w = \sum_i^N \alpha_i d_i x$$

$$\text{Therefore, } w^T w = \sum_i^N \alpha_i d_i w^T x \quad (2)$$

$$\sum_i^N \alpha_i d_i = 0 \quad (3)$$

Substituting (2) and (3) in (1)

$$w^T w + 0 - \sum_i^N \alpha_i = 0$$

$$w^T w = \sum_i^N \alpha_i$$

$$\|w\| = (\sum_i^N \alpha_i)^{1/2} \quad (4)$$

$$\text{Now, } \rho = \frac{2}{\|w\|}$$

$$\text{Therefore, } \rho = \frac{2}{(\sum_i^N \alpha_i)^{1/2}} \quad \text{Using, (4)}$$